

In [66]:

```

##imports from libraries
import pandas as pd
import numpy as np
import time
from sklearn import linear_model
import os
import pandas as pd

## LOAD INDIVIDUAL HOUSE POWER CONSUMPTION DATASET
data_power = pd.read_csv("household_power_consumption.txt", sep=";", low_m
data_power = data_power.to_numpy()
X_power = data_power[:,0:8]
Y_power = data_power[:,8]

##### LOAD GREENHOUSE GAS OBSERVING NETWORK DATASET
# since the documentation of the computer assinment is not clear, we take
# Since we have 15 tracers, we have 15 features and d=15. The last row of
file_list = os.listdir("ghg_data")
file_list = file_list[:5]
X_gas = np.zeros((2921,15))
Y_gas = np.zeros((2921))
i = 0
for file in file_list:
    df = pd.read_csv(f"ghg_data/{file}", sep="\s+", header=None, low_memo
    data_np_tmp = df.to_numpy()
    data_np_tmp = np.average(data_np_tmp, axis =1)
    X_gas[i,:] = data_np_tmp[:15]
    Y_gas[i] = data_np_tmp[15]
    i += 1

X_power = X_power.astype('float64')
Y_power = Y_power.astype('float64')
X_gas = X_gas.astype('float64')

n_power = 2075259
d_power = 8
lamda_power = 1/n_power

n_gas = 2921
d_gas = 15
lamda_gas = 1/n_gas

# start = time.time()

# time_power = time.time() - start
# tmp1 = np.matmul( np.transpose(X_power), Y_power )
# tmp2 = np.linalg.inv( (1/n_power)*( np.matmul(np.transpose(X_power), X_
# w_power = np.matmul(tmp2,tmp1)

start = time.time()
tmp1 = np.matmul( np.transpose(X_gas), Y_gas )
tmp2 = np.linalg.inv( (1/n_gas)*( np.matmul(np.transpose(X_gas), X_gas) +
w_gas = np.matmul(tmp2,tmp1)
time_gas = time.time() - start

# reg = linear_model.Ridge(alpha=0)

```

✓ (a) Completed in different file

9 (b) } code available  
9 (c) } but not running  
and no results  
uploaded

✗ (d) Not available  
(could not locate)

```
# start = time.time()
# ## Find the optimal linear regressor here:

# end = time.time()
```

```
-----
-
ValueError                                Traceback (most recent call last)
/var/folders/gh/5plzzznd6y17bcq3b6hhw6sw0000gp/T/ipykernel_79435/88924934
3.py in <module>
    32
    33
----> 34 X_power = X_power.astype('float64')
    35 Y_power = Y_power.astype('float64')
    36 X_gas = X_gas.astype('float64')

ValueError: could not convert string to float: '16/12/2006'
```

In [63]:

```
tmp1 = np.matmul( np.transpose(X_power), Y_power )
```

```
-----
-
TypeError                                Traceback (most recent call last)
/var/folders/gh/5plzzznd6y17bcq3b6hhw6sw0000gp/T/ipykernel_79435/303554249
8.py in <module>
----> 1 tmp1 = np.matmul( np.transpose(X_power), Y_power )

TypeError: can't multiply sequence by non-int of type 'float'
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: